

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.</small>				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE May 1998	3. REPORT TYPE AND DATES COVERED Technical - 98-03		
4. TITLE AND SUBTITLE Estimating the Parameters of the Linear Compartment Model		5. FUNDING NUMBERS DAAH04-96-1-0082		
6. AUTHOR(S) D. Kundu and A. Mitra				
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES) Center for Multivariate Analysis Department of Statistics 417 Thomas Bldg. Penn State University University Park, PA 16802		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211		10. SPONSORING / MONITORING AGENCY REPORT NUMBER ARO 35518.30-MA		
11. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.		12 b. DISTRIBUTION CODE		
13. ABSTRACT (Maximum 200 words) In this paper we consider the linear compartment model and consider the estimation procedures of the different parameters. We discuss a method to obtain the initial estimators, which can be used for any iterative procedures to obtain the least squares estimators. Four different types of confidence intervals have been discussed and they have been compared by computer simulations. We propose different methods to estimate the number of components of the linear compartment model. One data set has been used to see how the different methods work in practice. Reproduced From Best Available Copy				
14. SUBJECT TERMS Compartment Model, Confidence Intervals, Model Selection, Monte Carlo Simulations		15. NUMBER OF PAGES 21		
		16. PRICE CODE		
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

ESTIMATING THE PARAMETERS OF THE
LINEAR COMPARTMENT MODEL

Debasis Kundu and Amit Mitra

Technical Report 98-03

May 1998

Center for Multivariate Analysis
417 Thomas Building
Penn State University
University Park, PA 16802

19981228 108

Research work of authors was partially supported by the Army Research Office under Grant DAAHO4-96-1-0082. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

To appear in J. of Stat. Planning & Inf.

ESTIMATING THE PARAMETERS OF THE LINEAR COMPARTMENT MODEL

Debasis Kundu¹ and Amit Mitra²
Department of Mathematics
Indian Institute of Technology Kanpur
Pin 208016, India

ABSTRACT:

In this paper we consider the linear compartment model and consider the estimation procedures of the different parameters. We discuss a method to obtain the initial estimators, which can be used for any iterative procedures to obtain the least squares estimators. Four different types of confidence intervals have been discussed and they have been compared by computer simulations. We propose different methods to estimate the number of components of the linear compartment model. One data set has been used to see how the different methods work in practice.

Key Words and Phrases: Compartment Model, Confidence Intervals, Model Selection, Monte Carlo Simulations.

AMS Subject Classifications(1985): 62J02, 65D10

Short Running Title: Linear Compartment Model

e-mail: kundu@iitk.ernet.in

¹ Part of the work has been supported by a Grant (No. SR/OY/M-06/93) from the Department of Science and Technology, Government of India.

² Part of the work has been supported by the National Board of Higher Mathematics, Department of Atomic Energy, Government of India.

1. INTRODUCTION:

Fitting a sum of exponential model to equispaced data, popularly called as linear compartment model, to analyze a real data set is quite common to the scientists and engineers. See for example Anderson (1983), Seber and Wild (1989), Gallant (1987), Bates and Watts (1988) and the references there in. In this paper we consider the fitting of the following linear compartment model;

$$y(t) = \sum_{k=1}^p \alpha_k e^{\beta_k t} + \epsilon(t) \quad (1)$$

Here $y(t)$ is observed at the time points t_1, \dots, t_n , which are assumed to be equidistant. α 's and β 's are unknown parameters, p , is also assumed to be unknown. $\epsilon(t)$'s are error random variables, they are assumed to be independent and identically distributed (i.i.d.) with zero mean and finite variance σ^2 . The problem is, given a sample of size n , $y(t_1), \dots, y(t_n)$, estimate p , $\alpha_1, \dots, \alpha_p$ and also estimate β_1, \dots, β_p . Note that without loss of generality we can assume $\beta_1 < \beta_2 < \dots < \beta_p$. The ordering may be needed in the identification of the estimates. We also assume $p < \frac{n}{2}$. This assumption is realistic, infact in practice usually n is quite large compared to p .

This is a very old problem and well known problem in statistical literature. The first reference of this work goes back to the pionerring work of Prony in 1795. After that several articles have been appeared mainly to establish some iterative procedures to obtain the least squares estimators (LSE), when 'p' is known. See for example Barham and Drane (1972), Golub and Pereyra (1973), Osborne (1975, 1976), Osborne and Smyth (1986, 1994), Kahn et al. (1992). It is well known that this problem is a numerically difficult problem. Recently it is observed by Osborne and Smyth (1994) that the method originally proposed by Osborne (1975) can be very effective, provided the initial value is reasonably good. Infact it is observed that all the methods are quite sensitive to the initial value. Unfortunately no where in the literature except the work of Prony (1795) the choice of the initial value has been suggested. Infact we will see that in many cases the Prony's estimators can not be used as an initial guess. We discuss a method which can be used to obtain the initial guess of any one of the standard algorithm. We also propose four different types of the confidence intervals and compare their performances through Monte Carlo simulations. We also consider the estimation of 'p'. We propose to use the information theoretic criteria and the Cross Validation technique in this purpose. We apply these methods to one real life data set and see how the proposed methods work in practice.

The organisation of the paper is as follows: From Section 2 to Section 6, we assume that 'p' is known. In Section 2, we introduce the seperable regression technique in this case as it was originally proposed by Richards (1961). We introduce the Prony's estimator in Section 3 and the proposed initial value estimator is presented in Section 4. We discuss

different confidence intervals in Section 5 and some experimental results are presented in Section 6. Estimation of 'p' is considered in Section 7 and the data analysis of one data set is being done in Section 8. Finally we draw conclusions of our work in Section 9.

2. LEAST SQUARES ESTIMATORS:

Observe that under the assumption of additive i.i.d. error the LSE can be obtained by minimizing the residual sums of squares;

$$R(\alpha, \beta) = \sum_{t=t_1}^{t_n} \left[y(t) - \sum_{k=1}^p \alpha_k e^{\beta_k t} \right]^2 \quad (2)$$

with respect to α and β , where $\alpha = (\alpha_1, \dots, \alpha_p)$ and $\beta = (\beta_1, \dots, \beta_p)$. Here $R(\alpha, \beta)$ can be written in a matrix form as follows;

$$R(\alpha, \beta) = ((Y - A(\beta)\alpha)^T (Y - A(\beta)\alpha)) \quad (3)$$

where $Y = (y(t_1), \dots, y(t_n))$ and $A(\beta)$ is a $n \times p$ matrix of the following form

$$A(\beta) = \begin{pmatrix} e^{\beta_1 t_1} & \dots & e^{\beta_p t_1} \\ \vdots & \vdots & \vdots \\ e^{\beta_1 t_n} & \dots & e^{\beta_p t_n} \end{pmatrix}, \quad (4)$$

Now observe that for a fixed 'p', the LSE of α 's are

$$\hat{\alpha}(\beta) = [A^T(\beta)A(\beta)]^{-1}A^T(\beta)Y \quad (5)$$

Now if we replace $\hat{\alpha}(\beta)$ obtained from (5) in (3), we obtain;

$$R(\hat{\alpha}, \beta) = Q(\beta) = Y^T(I - P_A)Y \quad (6)$$

where $P_A = A(\beta)[A^T(\beta)A(\beta)]^{-1}A^T(\beta)$ is the projection matrix on the space spanned by the columns of $A(\beta)$. Therefore the LSE of (α, β) obtained by minimizing $R(\alpha, \beta)$ with respect to α, β is same as first obtaining the LSE, $\hat{\beta}$, of β by minimizing $Q(\beta)$ with respect to β and then obtain the LSE of α by putting $\hat{\beta}$ in place of β in (5). Osborne (1975) considered the minimization of $Q(\beta)$ with respect to β and he transformed the minimization problem to a non-linear eigenvalue problem and proposed an iterative scheme to solve it. Recently it is observed by Osborne and Smyth (1994) that the iterative scheme originally proposed by Osborne(1975) is numerically stable, i.e., if the initial value of the iterative procedure is close to the true value (within the radius of convergence) then the iterative procedure converges almost surely.

3. PRONY'S ESTIMATOR:

Prony observed the following (see for example Barrodale and Olesky; 1981). If

$$\mu_t = \sum_{k=1}^p \alpha_k e^{\beta_k t} \quad \text{for } t = t_1, \dots, t_n \quad (7)$$

then there exists $(p+1)$ constants, g_1, \dots, g_{p+1} , such that

$$g_1 \mu_{t_i} + g_2 \mu_{t_{i+1}} + \dots + g_{p+1} \mu_{t_{i+p}} = 0 \quad \text{for } i = 1, \dots, n. \quad (8)$$

Here $\mathbf{g} = (g_1, \dots, g_{p+1})$ is unique upto a constant multiplication and therefore without loss of generality we can assume $|\mathbf{g}| = \sum_{i=1}^{p+1} g_i^2 = 1$. Moreover it can be shown that if $h = t_2 - t_1$, then $e^{\beta_1 h}, \dots, e^{\beta_p h}$ are the roots of the following polynomial equation:

$$B(x) = g_1 + g_2 x + \dots + g_{p+1} x^{p+1} = 0 \quad (9)$$

Therefore there exists a one to one correspondence between \mathbf{g} , such that $|\mathbf{g}| = 1$ and $g_1 > 0$, with the nonlinear parameter β . Interestingly \mathbf{g} is independent of α (see Tufts and Kumaresan; 1982). Observe that (8) can be written in the following form;

$$\mathbf{M}\mathbf{g} = \mathbf{0} \quad (10)$$

where \mathbf{M} is the $n-p \times p+1$ matrix of μ 's. The rank of the matrix \mathbf{M} is p if $n-p > p+1$ and \mathbf{g} is the eigenvector corresponding to the zero eigenvalue of $\mathbf{M}^T \mathbf{M}$. This result can be used to obtain an initial estimator of β 's from the data. The idea is as follows. Assume that the error variance is small then form a matrix \mathbf{Y} from \mathbf{M} by replacing μ_t by y_t (assuming $E(y_t) = \mu_t$) for t_1, \dots, t_n . Then obtain $\hat{\mathbf{g}} = (\hat{g}_1, \dots, \hat{g}_p)$ as the eigenvector corresponding to the minimum eigenvalue of $\mathbf{Y}^T \mathbf{Y}$. Normalize $\hat{\mathbf{g}}$ s.t. $|\hat{\mathbf{g}}| = 1$ and $\hat{g}_1 > 0$. Construct the polynomial equation

$$\hat{B}(x) = \hat{g}_1 + \hat{g}_2 x + \dots + \hat{g}_{p+1} x^{p+1} = 0 \quad (11)$$

and obtain the roots of the polynomial (11) of the form $e^{\hat{\beta}_1 h}, \dots, e^{\hat{\beta}_p h}$. Note that we can always assume $\hat{\beta}_1 < \hat{\beta}_2 < \dots < \hat{\beta}_p$, because it is simply renaming the estimators and then the $\hat{\beta}_1, \dots, \hat{\beta}_p$ can be used as estimators of β_1, \dots, β_p . It usually behaves quite well as an initial guess provided the error variance is small. If the error variance is large then there are usually two kinds of problems. It may be possible that some roots of the polynomial equation (11) are complex or some roots may be negative. In both the cases it is not possible to obtain the estimators $\hat{\beta}_1, \dots, \hat{\beta}_p$. Osborne (1975) used these estimators as the initial value for his numerical experiments. In our numerical experiments we observed that Prony's estimator does not exist in some situations.

4. MODIFIED PRONY ESTIMATOR:

Prony's idea has been used extensively in the Signal Processing literature to estimate the frequencies of an undamped exponential model. See for example the work of Ulrych

and Clayton (1976), Kundu (1994, 1995), Kundu and Mitra (1995a, 1995b), Tufts and Kumaresan (1982). Several modifications have been suggested by the above authors for the complex valued signals. Unfortunately these methods can not be applied in the case of real data sets. In this section we propose a new estimator which is a modification of the existing Prony's estimator, and it can be used quite effectively as an initial estimator for different algorithms.

First consider the same model as (1) and y_t and μ_t are defined same as in Section 3. Now choose L , such that, $p < L < n - p$. It can be shown similarly as Tufts and Kumaresan (1982) that there exists b_0, \dots, b_L such that

$$b_0\mu_{t_i} + b_1\mu_{t_{i+1}} + \dots + b_L\mu_{t_{i+L}} = 0 \quad \text{for } i = 1, \dots, n - L. \quad (12)$$

where $\sum_{i=1}^L b_i^2 = 1$ and $e^{\beta_1 h}, \dots, e^{\beta_p h}$ are the p roots, among the L roots, of the following polynomial equation;

$$b_0 + b_1x + \dots + b_Lx^L = 0 \quad (13)$$

Now we can write (12) in the matrix form as;

$$\mathbf{A}\mathbf{b} = \mathbf{0} \quad (14)$$

Since the rank of the matrix \mathbf{A} is p , therefore \mathbf{b} is not unique as long as $L > p$ and $L + 1 < n - L$. Observe that (14) is equivalent to

$$\mathbf{A}^T \mathbf{A} \mathbf{b} = \mathbf{0} \quad (15)$$

Consider the spectral decomposition of $\mathbf{A}^T \mathbf{A}$ as follows;

$$\mathbf{A}^T \mathbf{A} = \sum_{i=1}^{L+1} \lambda_i \mathbf{a}_i \mathbf{a}_i^T \quad (16)$$

where $\lambda_1 > \lambda_2 > \dots > \lambda_{L+1}$ and \mathbf{a}_i 's are the orthonormal eigenvectors corresponding to λ_i . Since the rank of $\mathbf{A}^T \mathbf{A}$ is p , this implies $\lambda_{p+1} = \dots = \lambda_{L+1} = 0$. Therefore any vector \mathbf{b} , which belongs to the linear space spanned by $[\mathbf{a}_{p+1}, \dots, \mathbf{a}_{L+1}]$ will satisfy (16). Let's consider the null space of $\mathbf{A}^T \mathbf{A}$ as follows;

$$\mathbf{N} = [\mathbf{a}_{p+1}^T : \dots : \mathbf{a}_{L+1}^T] \quad (17)$$

Now since the rank of \mathbf{N} is $L - p + 1$, from (8) it follows that there are $(L - p + 1)$ independent vectors of the following form;

$$\mathbf{N}_g = \begin{bmatrix} g_1 & 0 & \dots & 0 \\ \cdot & g_1 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ g_{p+1} & g_p & \dots & 0 \\ 0 & g_{p+1} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & g_1 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & g_{p+1} \end{bmatrix} \quad (18)$$

such that the space spanned by \mathbf{N} and \mathbf{N}_g are equal. This idea can be used to estimate \mathbf{g} from y_t 's as follows.

Replace μ_t 's by y_t 's to form the matrix $\hat{\mathbf{A}}$, obtain the spectral decomposition of $\hat{\mathbf{A}}^T \hat{\mathbf{A}}$ as follows;

$$\hat{\mathbf{A}}^T \hat{\mathbf{A}} = \sum_{i=1}^{L+1} \hat{\lambda}_i \hat{\mathbf{a}}_i \hat{\mathbf{a}}_i^T; \quad (19)$$

where $\hat{\lambda}_1 > \dots > \hat{\lambda}_L$ and $\hat{\mathbf{a}}_i$ are the orthonormal eigenvectors corresponding to $\hat{\lambda}_i$. Although the rank of $\mathbf{A}^T \mathbf{A}$ is p , but the rank of $\hat{\mathbf{A}}^T \hat{\mathbf{A}}$ is $L + 1$. Obtain the estimation of the null space of \mathbf{N} as;

$$\hat{\mathbf{N}} = [\hat{\mathbf{a}}_{p+1} : \dots : \hat{\mathbf{a}}_{L+1}] \quad (20)$$

Now we would like to obtain a basis of $\hat{\mathbf{N}}$ of the form (18), which can be used to estimate \mathbf{g} . Let's write

$$\hat{\mathbf{N}} = \begin{bmatrix} \hat{\mathbf{N}}_1 \\ \hat{\mathbf{N}}_2 \end{bmatrix} \quad (21)$$

where $\hat{\mathbf{N}}_1$ is a $p + 1 \times L + 1 - p$ matrix and $\hat{\mathbf{N}}_2$ is a $L - p \times L + 1 - p$ matrix. Since $\hat{\mathbf{N}}_2$ is of rank $L - p$, therefore there exists an unique vector \mathbf{b}^1 upto a constant multiplication such that $\hat{\mathbf{N}}_2 \mathbf{b}^1 = \mathbf{0}$. Therefore if we denote $\hat{\mathbf{g}}^1 = \hat{\mathbf{N}}_1 \mathbf{b}^1$, then;

$$\hat{\mathbf{N}} \mathbf{b}^1 = \begin{bmatrix} \hat{\mathbf{g}}^1 \\ \mathbf{0} \end{bmatrix} \quad (22)$$

Proceeding in the same manner, we can say that there exists vectors $\mathbf{b}^1, \dots, \mathbf{b}^{L-p+1}$,

such that;

$$\hat{N} [b^1 : \dots : b^{L-p+1}] = \begin{bmatrix} \hat{g}^1 & 0 & \dots & 0 \\ 0 & \hat{g}^2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \hat{g}^{L-p+1} \end{bmatrix} \quad (23)$$

Now any one or all or a subset of \hat{g}^k , for $k = 1, \dots, L - p + 1$, can be used to estimate g . It is important to note that when $E(y_i) = y_i$, for $i = 1, \dots, n$, then $\hat{g}^k = g$, for $k = 1, \dots, L - p + 1$. Now we choose that particular \hat{g}^k from the $L - p + 1$, \hat{g} 's so that it 'fits' the data best, i.e. for each \hat{g}^k ; $1 \leq k \leq L - p + 1$, we obtain the estimates of $(\alpha_1, \dots, \alpha_p)$ and $(\beta_1, \dots, \beta_p)$ and then calculate the residual sums of squares. We choose that set of α 's and β 's for which the residual sums of squares is minimum. Although we may get some infisible roots with respect to some \hat{g}^k , but it is observed that not all of them will give infisible roots.

The method can be easily applied to the following model;

$$y_t = \alpha_0 + \sum_{k=1}^p \alpha_k e^{\beta_k t} + \epsilon_t \quad (24)$$

that is when the constant term is also present. We apply the above procedure to the data set $y_t - \bar{y}$, when $\bar{y} = \sum_{i=1}^n \frac{y_i}{n}$ if it is known that all the β 's are negative. Otherwise we apply the above method to the complete model namely

$$y_t = \sum_{k=0}^p \alpha_k e^{\beta_k t} + \epsilon_t \quad (25)$$

and then put $\hat{\beta}_0 = 0$.

Another important aspect is the choice of L . Although in theory L can be any integer satisfying $p < L < n - p$ but it is observed that the performance of the modified Prony estimator depends on L . It is observed (numerically) that as L increases at the beginning the performance of the modified Prony estimator increases in terms of the mean squared errors but after that it starts decreasing. It is observed that for $L \approx \frac{n}{3}$ the performance is the best, although no theoretical justifications can be provided. It seems more work is needed in this direction.

5. CONFIDENCE INTERVALS:

In this section we propose three different confidence intervals other than the classical one as is available in the literature (Gallant; 1987, page 105).

One important aspect of this model is its asymptotic properties, namely the consistency and the asymptotic normality. If we take a particular case of this model, namely $t_i = i$, for $i = 1, \dots, n$ and $p = 1, \alpha_1 = 1$, then it follows from Wu (1981) that any estimator of β is inconsistent. Infact to obtain the consistency and the asymptotic normality properties we have to make the boundedness assumptions on t_i 's and we have to rewrite the model as that of Kundu (1994b). Under the same set of assumptions as of Kundu (1994b), we have the following result.

If we denote $\theta = (\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_p)$, and $\hat{\theta}$ to be the LSE of θ of the model (1) and θ_0 be the true parameter value of θ in (1), then we can say;

Theorem 1 The LSE $\hat{\theta}$ of θ of the model (1) is strongly consistent and

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightarrow N(0, \sigma^2 \mathbf{A}^{-1}) \quad (26)$$

where σ^2 is the error variance, $\mathbf{A} = ((a_{ij}))$, with

$$a_{ij} = \frac{1}{b-a} \int_a^b h'_i(\theta, t) h'_j(\theta, t) dt \quad (27)$$

where

$$h'_i(\theta, t) = \frac{\delta}{\delta \theta_i} h(\theta, t), \quad \text{and} \quad h(\theta, t) = \sum_{k=1}^p \alpha_k e^{\beta_k t} \quad (28)$$

Here a and b are the lower and the upper bound of the sampling times. In our case we take $a = t_1$ and $b = t_n$. Now we can use (26) to obtain $100(1 - \alpha)\%$ confidence interval of θ . Interested readers may refer to Kundu and Mitra (1996) for a detailed proof of the above theorem.

We also propose two Bootstrap confidence intervals. One is the percentile Bootstrap and the other one is the Bootstrap-t confidence interval. Percentile Bootstrap confidence interval can be described as follows:

[1] From the sample y_t ; $t = t_1, \dots, t_n$, estimate α 's and β 's.

[2] Estimate $\hat{\epsilon}_t = y_t - \hat{y}_t$, where

$$\hat{y}_t = \sum_{j=1}^p \hat{\alpha}_j e^{\hat{\beta}_j t}$$

for $t = t_1, \dots, t_n$.

[3] Draw a random sample of size n from $\{\hat{\epsilon}_{t_1}, \dots, \hat{\epsilon}_{t_n}\}$, let it be $\{\hat{\epsilon}_{B_1}, \dots, \hat{\epsilon}_{B_n}\}$.

[4] Obtain the Bootstrap sample $\hat{y}_{t_1}^*, \dots, \hat{y}_{t_n}^*$; where

$$y_{t_i}^* = \hat{y}_{t_i} + \hat{\epsilon}_{t_i} \quad \text{for} \quad i = 1, \dots, n$$

- [5] Estimate α 's and β 's from y_t^* 's and denote it by α^{**} 's and β^{**} 's.
- [6] Repeat the process (3) to (5) NBOOT times.
- [7] From the NBOOT estimators obtain the lower and the upper bound of the $100(1 - \alpha)\%$ percentile Bootstrap confidence interval for each α_k and β_k .

We propose the following algorithm for computing the Bootstrap-t confidence intervals.

- [1] From the sample y_t ; $t = t_1, \dots, t_n$, estimate α 's and β 's.
- [2] Estimate $\hat{\epsilon}_t = y_t - \hat{y}_t$, where

$$\hat{y}_t = \sum_{j=1}^p \hat{\alpha}_j e^{\hat{\beta}_j t}$$

for $t = t_1, \dots, t_n$.

- [3] Draw a random sample of size n from $\{\hat{\epsilon}_{t_1}, \dots, \hat{\epsilon}_{t_n}\}$, let it be $\{\hat{\epsilon}_{B_1}, \dots, \hat{\epsilon}_{B_n}\}$.
- [5] Estimate α 's and β 's from y_t^* and also estimate σ^2 as $\hat{\sigma}_B^2$ by mean residual sums of squares.
- [6] For each α or β (say θ), compute

$$T = \frac{\sqrt{n}(\hat{\theta}^* - \hat{\theta})}{\hat{\sigma}_B}$$

- [7] Repeat the steps (3) to (6) NBOOT times.
- [8] From the NBOOT estimators, obtain the lower and upper bound of $100(1 - \alpha)\%$ Bootstrap-t confidence intervals for each α_k and β_k .

In the next section we have performed some numerical experiments to see how the different methods behave for finite sample.

6. NUMERICAL EXPERIMENTS AND DISCUSSIONS:

In this section we perform some numerical experiments mainly to see how the Modified Prony's estimator behaves compared to the usual Prony estimator as an initial guess for different iterative procedures. We compare Osborne's method, which is well known to work very good for this problem, with that of the usual Gauss Newton method. We also compare different confidence intervals with respect to their coverage probability and average length. We consider the following model;

$$y_t = -6.0e^{-.232t} + 3.0e^{.0119t} + \epsilon_t; \quad t = 1, \dots, n \quad (29)$$

ϵ_t is taken to be i.i.d., normal random variables with mean zero and finite variance. We take the values of n as 25, 50, 75 and 100 and the different values of σ are .01, .05 and .1. For each n and σ , we generate 500 samples and compute the Prony estimators (PE) and the modified Prony estimators (MPE). Then using the PE and MPE as the initial guesses we estimate α 's and β 's using Gauss Newton (GN) and the Osborne's (1975) (OS) method. We compute the mean squared errors (MSE) of the different estimators. We also report the median number of iteration counts and the maximum number of iteration taken over 500 replications. We also report the the number of times the PE does not give the feasible estimators.

We also compute the confidence intervals of the different parameters by four different methods namely the classical one , the proposed one and the two Bootstrap confidence intervals. We compute the average length and the coverage probability. We take the nominal level to be 90% in our experiments. We report all the results in Tables: 1-9. In Table 1-4, we report the average estimates and the MSE's of the PE, MPE and the LSE. In each box, the first row, the second row and the third row indicate the results of the PE, MPE and LSE respectively. In each row the first figure represents the average value of the corresponding estimator and in the bracket we present the corresponding MSE. Table 5 represents the comparison of the OS method with that of the GN method when the starting value is different. We also present the median number of iteration count and the maximum iteration taken for both OS method and GN method, when the initial guess is the usual PE and also when the initial guess is the MPE. In each box, the first figure represents the median iteration count and in bracket the figure indicates the maximum number of iterations required over 500 replications for GN method. Similarly the second figure represents the result for OS method. In Tables 6-9, we present the comparison of the performances of the four different confidence intervals. In each table we present the average length over 500 replications and in bracket we present the coverage probability.

From the Tables 1-4, it is observed that as sample size increases or the error variance decreases the MSE of all the estimators, namely PE, MPE and LSE, decrease. It is clear that MPE behave better than PE and the LSE behave better than MPE in all the cases in terms of lower MSE. In our simulations it is observed that although PE does not exist in certain cases but MPE always exist. It is observed that PE may not exist if the sample size is small and the error variance is large. If the sample size is large or the error variance is small PE exist. From Table 5, it is observed that as sample size increases or the error variance decreases, the median number of iteration count or the maximum number of iteration count decreases in all the cases. It is also observed that the required number of iteration is much less if the MPE are used as the initial guess to obtain the LSE and it is very prominent if the error variance is high. It is also observed that between the OS method and the GN method the OS method takes less number of iterations compared to GN method in all the cases considered. It indicates that it is better to use the OS method

rather than the GN method in general. Although if we use the MPE as the initial guesses, it does not make much differences.

About the different confidence intervals, it is observed that they behave quite differently in terms of the average confidence length and the coverage probability. It is observed that as sample size increases or the error variance decreases the average length of the confidence intervals decreases in all the four cases. It is also observed that as sample size increases the coverage probability becomes closer to the nominal value. Among the different methods it is observed that the classical method can not maintain the nominal coverage probability in many situations, where as the asymptotic method or the Bootstrap procedure behaves better than the classical method in the sense that the coverage probability is higher than that of the classical one. Although in certain situations it also can not maintain the nominal coverage probability and in certain situations the coverage probability is much higher than the nominal level. Bootstrap-t method works better in that sense, as all the cases considered it is capable of maintaining the nominal level although the average length is higher in most of the cases. It is also computer intensive. Comparing all these we recommend to use Bootstrap-t and if we want to avoid heavy computation, we may use the asymptotic confidence interval.

7. ESTIMATION OF p :

In this section we propose different methods to estimate p . We assume that $p \leq M$, where M is some fixed unknown integer. We propose to use Akaike Information Theoretic Criteria (AIC) and Bayes Information Theoretic Criteria in this setup as it was originally proposed by Akaike (1971), Schwartz (1978) and Risannen (1978) in the general model selection setup. See Kundu and Murali (1996) for the comparison of the different methods. Rao (1986) also suggested to use different information theoretic criteria to estimate the number of signals in a similar problem in Signal Processing. See for example Kundu (1993) and Mitra and Kundu (1996) in this connection. AIC and BIC take the following form in this case:

$$AIC(k) = n \ln \hat{R}_k^2 + 2 \times \text{number of free parameters} \quad (30)$$

$$BIC(k) = n \ln \hat{R}_k^2 + \frac{1}{2} \ln n \times \text{number of free parameters} \quad (31)$$

AIC (BIC) chooses that k for which (30) [(31)] is minimum. Here \hat{R}_k^2 is the residual sums of squares when, the model order is k , i.e.

$$\hat{R}_k^2 = \sum_{t=1}^n (y_t - \sum_{i=1}^k \hat{\alpha}_i e^{\hat{\beta}_i t})^2 \quad (32)$$

We can use the Cross Validation technique also to estimate p . We propose the following procedure for the Cross Validation.

- [1] Take $p = 1$.
- [2] From the data y_{t_1}, \dots, y_{t_n} , delete y_{t_k} ; $k = 1, \dots, n$.
- [3] Estimate α 's and β 's from $y_{t_1}, \dots, y_{t_{k-1}}, y_{t_{k+1}}, \dots, y_{t_n}$ by using the missing value technique of Kundu and Kundu (1994).
- [4] Estimate y_{t_k} by \hat{y}_{t_k} and obtain the Cross Validatory error

$$\sum_{k=1}^n (y_{t_k} - \hat{y}_{t_k})^2.$$

- [5] Repeat (2) to (4) for different values of $p = 1, 2, \dots, M$, and choose that value of 'p' for which the Cross Validatory error is minimum.

In the next section we use these techniques to determine the order of the model for one real life data set.

8. DATA ANALYSIS:

In this section we consider one real life data set from Osborne (1972, page 185). The data are reported in the following form (t, y_t) :

(0,.844), (10,.908), (20,.932), (30,.936), (40,.925), (50,.908), (60,.881), (70,.850), (80,.818), (90,.784), (100,.751), (110,.718), (120,.685), (130,.658), (140,.628), (150,.603), (160,.580), (170,.558), (180,.538), (190,.522), (200,.506), (210,.490), (220,.478), (230,.467), (240,.457), (250,.448), (260,.438), (270,.431), (280,.424), (290,.420), (300,.414), (310,.411), (320,.406).

Observe that in this case $n = 33$, $h = 10.0$, $a = t_1 = 0.0$ and $b = t_{33} = 320$. From the plot it is clear that the sum of exponentials can be tried to fit the data. First we use different order determination criteria as it is proposed in Section 7. Different information theoretic criteria take the following form:

Model	R_k^2	NOP	AIC	BIC
α_0	1.1529	2	8.6952	8.1918
$\alpha_1 e^{\beta_1 t}$.05105	3	-92.1733	-92.9286
$\alpha_0 + \alpha_1 e^{\beta_1 t}$.05089	4	-90.2769	-91.2839
$\sum_{k=1}^2 \alpha_k e^{\beta_k t}$.05082	5	-88.3223	-89.5810
$\alpha_0 + \sum_{k=1}^2 \alpha_k e^{\beta_k t}$.553E-4	6	-311.4900	-313.0008
$\sum_{k=1}^3 \alpha_k e^{\beta_k t}$.550E-4	7	-309.6699	-311.4320
$\alpha_0 + \sum_{k=1}^3 \alpha_k e^{\beta_k t}$.525E-4	8	-309.0169	-311.0310

Therefore it is observed from the above table that AIC and BIC choose the following model:

$$y_t = \alpha_0 + \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t} + \epsilon_t \quad (33)$$

Observe that Osborne (1975) and Osborne and Smyth (1995) also use the same model to fit this data. They did not mention why they have used that model. Osborne used the initial estimates as $\alpha_0 = .5$, $\alpha_1 = 1.5$, $\alpha_2 = -1.0$, $\beta_1 = -.01$, $\beta_2 = -.02$. He did not mention how he had obtained those initial estimates. With those initial estimates he obtained the following least squares estimates: $\alpha_0 = .3754$, $\alpha_1 = 1.9358$, $\alpha_2 = -1.4647$, $\beta_1 = -.01287$, $\beta_2 = -.02212$ after 27 iterations. We obtain the Modified Prony estimates with $L = 11$, as follows in this case: $\alpha_0 = .1993$, $\alpha_1 = 1.3670$, $\alpha_2 = -.7157$, $\beta_1 = -.0087$, $\beta_2 = -.0261$. Using these values as the initial estimates the usual Gauss Newton algorithm converges after 5 iterations to the same value as that of Osborne. The Osborne's method takes 4 iterations to converge in this case. The original data and the fitted curve are shown in Figure 1a. The observed and the fitted values are shown in Figure 1b. We obtain the following residuals:

-2.499E-03, 4.611E-03, 1.171E-03, -8.564E-04, -2.649E-03, 7.331E-05, -2.650E-04,
 -3.521E-04, 8.159E-04, 7.785E-04, 1.487E-03, 1.209E-03, -5.484E-04, 1.902E-03,
 -6.150E-04, -1.783E-04, 2.069E-04, -4.138E-04, -9.608E-04, 6.685E-04, 5.900E-04,
 -1.073E-03, -1.975E-04, 3.402E-04, 6.586E-04, 8.703E-04, -9.186E-04, -6.096E-04,
 -1.111E-03, 6.592E-04, -2.205E-04, 1.318E-03, 3.384E-04.

We perform run test (Draper and Smith; 1981, page 159) and Durbin-Watson test (Draper and Smith; 1981, page 162) on the residuals. It is observed that $z = .3594$ and $d = 2.0437$ for run test and Durbin-Watson test respectively. Therefore both the tests confirm the independent assumptions of the error random variables.

9. CONCLUSIONS:

In this paper we consider the linear compartment model as it is defined in (1). We propose MPE to obtain the initial estimators, which is a very important problem. It is observed that the proposed initial estimator can be used for any standard algorithm to obtain the LSE's. We observe that although PE may not exist in certain situation but MPE always exist. We compare the usual GN method with that of the OS method and it is observed that the OS method works very well compared to that of GN if the initial guess is PE but if we use the MPE as the initial guesses it does not make much difference, i.e. we donot need any special purpose algorithm to solve this particular problem. We also propose different confidence intervals and it is observed that the Bootstrap-t works the best in terms of the required coverage probability. We consider the problem of estimating

the number of terms in the linear compartment model. We propose to use AIC or BIC to estimate the number of terms. We also provide the scheme to estimate the number of terms by the Cross Validation technique. It seems any one of them can be used. Since the Cross Validation is more computer intensive AIC or BIC may be preferred.

Acknowledgements: The authors would like to thank one referee for his constructive suggestions on the earlier draft of this paper.

REFERENCES:

- [1] Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", *Proc. 2nd. International Symposium on Information Theory*, Academia Kiado, 267-281.
- [2] Anderson, D.H. (1983), **Compartmental Modelling and Tracer Kinetics**, Lecture Notes in Biomathematics, Springer Verlag, New York.
- [3] Bates, D.M. and Watts, D.G. (1988), **Nonlinear Regression Analysis and its Applications**, John Wiley, New York.
- [4] Barham, R.H. and Drane, W. (1972), "An algorithm for least squares estimation of non-linear parameters when some of the parameters are linear", *Technometrics*, 14, 3, 757-766.
- [5] Barrodale, I and Oleski, D.D. (1981), "Exponential approximation using Prony's method", **The Numerical Solution of Nonlinear Problems**, Eds. Baker, C.T.H. and Phillips, C., 258-269.
- [6] Draper, N.R. and Smith, H. (1981), **Applied Regression Analysis**, 2nd. ed. Wiley, New York.
- [7] Gallant, A.R. (1987), **Nonlinear Statistical Models**, John Wiley, New York.
- [8] Golub, G.H. and Pereyra, V. (1973), "The differentiation of pseudo- inverses and non-linear least squares problems whose variables separate", *SIAM Jour. Num. Anal.*, 10, 2, 413-432.
- [9] Kahn, M., Mackisack, M.S., Osborne, M.R. and Smyth, G.K. (1992), "On the consistency of Prony's method and related algorithms", *Jour. Comp. and Graph. Stat.*, 1,4, 329-349.
- [10] Kundu, D. (1992), "Detecting the number of signals for undamped exponential signals using information theoretic criteria", *Jour. Stat. Comp. Simul.*, 44, 117-131.

- [11] Kundu, D. (1995), "Small sample properties of some parametric and non parametric methods for detection of signals by Monte Carlo simulations", *IEE Proc. Vision Image and Signal Processing*, 142, 3, 181-186.
- [12] Kundu, D. (1994a), "Estimating the parameters of complex valued exponential signals", *Comp. Stat. and Data Anal.*, 18, 525-534.
- [13] Kundu, D. (1994b), "A modified Prony algorithm for sum of damped or undamped exponential signals", *Sankhya, A*, 56, 3, 525-544.
- [14] Kundu, D. and Kundu, R. (1995), "Consistent estimates of superimposed exponential signals when some observations are missing", *Jour. Stat. Plann. and Inf.*, 44, 205-218.
- [15] Kundu, D. and Mitra, A. (1995a), "Consistent method of estimating the superimposed exponential signals", *Scand. Jour. Stat.*, 22, 1, 73-82.
- [16] Kundu, D. and Mitra, A. (1995b), "Estimating the parameters of Damped exponential signals; A non iterative approach", *Signal Processing*, 46, 3, 363-368.
- [17] Kundu, D. and Mitra, A. (1996), "Fitting a sum of exponentials to equispaced data", Technical Report, I.I.T. Kanpur.
- [18] Kundu, D. and Murali, G. (1996), "Model selection in linear regression", *Comp. Stat. Data Anal.*, 22, 461-469.
- [19] Mitra, A. and Kundu, D. (1996), "Estimating the number of signals for the damped exponential model", Technical Rreport, I.I.T. Kanpur.
- [20] Osborne, M. R. (1972), "Some aspects of non-linear least squares calculations", **Numerical Methods for Nonlinear Optimization**, Ed. F.A. Lootsma, Academic Press, New York.
- [21] Osborne, M.R. (1975), "Some special nonlinear least squares problems", *SIAM Jour. Num. Anal.*, 12, 571-592.
- [22] Osborne, M.R. (1976), "Nonlinear least squares, the Levenberg algorithm revisited", *Jour. Austr. Math. Soc.*, B, 19, 343-357.
- [23] Osborne, M.R. and Smyth, G.K. (1986), "An algorithm for exponential fitting revisited", *Jour. App. Prob.*, 418-430.
- [24] Osborne, M.R. and Smyth, G. (1994), "A modified Prony algorithm for fitting sum of exponential functions", *SIAM Jour. Comp*, 16,1, 119-138.

- [25] Prony, R. (1795), "Essai experimentale et analytique", *Jour. Ecole Polytechnique (Paris)*, 1, 24-76.
- [26] Rao, C.R. (1988), "Some results in signal detection", *Statistical Decision Theory and Related Topics IV*, Eds. Gupta, S.S. and Berger, J.O., 2, 319-332, Springer Verlag, New York.
- [27] Richards, F.S.G. (1961), "A method of maximum likelihood estimation", *Jour. Royal Stat. Soc.*, B, 23, 469-475.
- [28] Rissanen, J. (1978), "Modelling by shortest data description", *Automatica*, 14, 465-471.
- [29] Schwartz, G. (1978), "Estimating the dimension of a model", *Ann. Stat.*, 6, 461-464.
- [30] Seber, A. and Wild, B. (1989), **Nonlinear Regression**, John Wiley, New York.
- [31] Tufts, D.W. and Kumaresan, R. (1982), "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood", *IEEE Special issue on Spectral Estimation*, 70, 975-989.
- [32] Ulrych, T.J. and Clayton, R.W. (1976), "Time series modelling and maximum entropy", *Phys. Earth and Planetary Interiors*, 12, 188-200.
- [33] Wu, C.F.J. (1981), "Asymptotic theory of nonlinear least squares estimation", *Ann. Math. Stat.*, 9, 501-513.

Table 1
Average Estimates and the Mean Squared Errors of the Different Estimators
N = 25.

σ	α_1	α_2	β_1	β_2
.01	-6.008(.0071)	2.999(.0087)	-.2323(6.96E-5)	.0119(2.16E-6)
	-6.005(.0037)	2.999(.0064)	-.2321(5.04E-5)	.0119(1.53E-6)
	-6.000(.0002)	3.000(.0003)	-.2320(2.62E-6)	.0119(7.34E-8)
.05	-6.172(.2921)	3.135(.3520)	-.2354(1.83E-3)	.0114(6.22E-5)
	-6.120(.1308)	3.044(.1980)	-.2349(1.34E-3)	.0117(3.83E-5)
	-6.005(.0065)	3.002(.0083)	-.2324(6.49E-5)	.0119(1.86E-6)
.1*	-6.854(1.321)	3.297(2.321)	-.2821(1.23E-2)	.0120(6.78E-4)
	-6.321(.9316)	3.147(1.231)	-.2524(5.37E-3)	.0121(1.40E-4)
	-6.020(.0290)	3.012(.0366)	-.2329(2.59E-4)	.0118(7.69E-6)

* The ordinary Prony Estimator did not exist 25 times, so the results of the Prony Estimators are based on the average of 475 replications.

Table 2
Average Estimates and the Mean Squared Errors of the Different Estimators
N = 50.

σ	α_1	α_2	β_1	β_2
.01	-6.006(.0070)	3.002(.0025)	-.2322(5.08E-5)	.0119(2.19E-7)
	-6.000(.0030)	2.999(.0006)	-.2321(2.57E-5)	.0119(4.79E-8)
	-6.000(1.49E-4)	3.000(2.44E-5)	-.2320(8.97E-7)	.0119(1.96E-9)
.05	-6.118(.2038)	3.069(.0695)	-.2354(1.31E-3)	.0117(5.78E-6)
	-6.034(.0637)	2.977(.0182)	-.2383(7.14E-4)	.0121(1.47E-6)
	-6.001(.0038)	3.000(6.07E-4)	-.2322(2.26E-5)	.0119(4.91E-8)
.1*	-6.304(1.179)	3.236(.7069)	-.2365(5.62E-3)	.0109(3.34E-5)
	-6.168(.3166)	2.970(.0671)	-.2510(3.05E-3)	.0122(5.14E-6)
	-6.003(.0151)	3.000(.0024)	-.2325(9.16E-5)	.0119(1.96E-7)

* The ordinary Prony Estimator did not exist 11 times, so the results of the Prony Estimators are based on the average of 489 replications.

Table 3
Average Estimates and the Mean Squared Errors of the Different Estimators
N = 75

σ	α_1	α_2	β_1	β_2
.01	-5.998(7.18E-3)	2.998(1.91E-3)	-.2323(4.80E-5)	.0119(7.82E-8)
	-6.003(2.53E-3)	2.999(2.68E-4)	-.2324(2.30E-5)	.0119(1.06E-8)
	-6.000(1.49E-4)	3.000(6.47E-6)	-.2320(4.91E-7)	.0119(2.38E-10)
.05	-6.071(.2001)	3.041(.0532)	-.2356(1.25E-3)	.0119(2.10E-6)
	-6.055(.0571)	2.991(6.55E-3)	-.2373(5.53E-4)	.0119(2.60E-7)
	-6.000(.0037)	3.000(1.60E-4)	-.2322(1.24E-5)	.0119(5.92E-9)
.1	-6.157(1.127)	3.125(.3311)	-.2370(5.45E-3)	.0117(1.10E-5)
	-6.244(.2670)	2.986(.0241)	-.2521(2.42E-3)	.0120(9.55E-7)
	-6.002(.0151)	3.000(6.34E-4)	-.2324(5.03E-5)	.0119(2.35E-8)

Table 4
Average Estimates and the Mean Squared Errors of the Different Estimators
N = 100

σ	α_1	α_2	β_1	β_2
.01	-5.998(7.07E-3)	2.998(1.15E-3)	-.2323(4.63E-5)	.0119(2.53E-8)
	-6.001(2.11E-3)	3.000(1.40E-4)	-.2326(2.23E-5)	.0119(2.94E-9)
	-6.000(1.49E-4)	3.000(3.77E-6)	-.2320(4.43E-7)	.0119(7.12E-11)
.05	-6.059(.1959)	3.030(.0305)	-.2360(1.21E-3)	.0120(6.61E-7)
	-6.102(.0411)	3.004(3.72E-3)	-.2389(5.48E-4)	.0118(7.94E-8)
	-6.000(.0033)	3.000(9.41E-5)	-.2322(1.12E-5)	.0119(1.77E-9)
.1	-6.103(1.115)	3.074(.2213)	-.2386(5.21E-3)	.0119(3.68E-6)
	-6.349(.2121)	3.004(.0129)	-.2577(1.96E-3)	.0119(2.70E-7)
	-6.002(.0152)	3.000(3.75E-4)	-.2324(4.53E-5)	.0119(7.08E-9)

Table 5
Comparison of the Osborne's Method and the Gauss Newton Method

N	σ	PRONY	MODIFIED PRONY
25	.01	5(8), 5(7)	2(3), 2(3)
	.05	7(9), 6(9)	3(9), 3(9)
	.1	15(100), 12(88)	4(41), 4(36)
50	.01	4(8), 4(6)	2(2), 2(2)
	.05	6(9), 5(9)	3(3), 3(3)
	.1	12(70), 10(70)	3(5), 3(4)
75	.01	4(6), 4(6)	2(2), 2(2)
	.05	4(9), 4(9)	2(3), 2(3)
	.1	10(63), 8(63)	3(4), 3(4)
100	.01	3(3), 3(3)	2(2), 2(2)
	.05	4(5), 4(5)	2(3), 2(3)
	.1	7(19), 6(15)	3(4), 3(4)

Table 6
Different Confidence Intervals
N = 25

σ	Methods	α_1	α_2	β_1	β_2
.01	Asymptotic	.0527(.87)	.0600(.92)	.0057(.93)	.0009(.92)
	Classical	.0494(.83)	.0564(.91)	.0051(.90)	.0009(.91)
	Bootstrap	.0497(.84)	.0563(.92)	.0051(.91)	.0009(.89)
	Bootstrap-t	.0569(.91)	.0648(.92)	.0058(.93)	.0009(.93)
.05	Asymptotic	.2661(.87)	.3012(.91)	.0285(.94)	.0005(.91)
	Classical	.2489(.82)	.2827(.91)	.0256(.90)	.0004(.91)
	Bootstrap	.2520(.85)	.2845(.91)	.0253(.91)	.0004(.89)
	Bootstrap-t	.2895(.91)	.3275(.92)	.0290(.93)	.0005(.93)
.1	Asymptotic	.5476(.89)	.6133(.92)	.0571(.93)	.0092(.91)
	Classical	.5106(.83)	.5754(.91)	.0512(.90)	.0086(.91)
	Bootstrap	.5462(.85)	.6234(.92)	.0527(.93)	.0091(.92)
	Bootstrap-t	.6290(.90)	.7124(.92)	.0604(.92)	.0104(.92)

Table 7
Different Confidence Intervals
N = 50

σ	Methods	α_1	α_2	β_1	β_2
.01	Asymptotic	.0493(.92)	.0148(.92)	.0030(.88)	.0001(.88)
	Classical	.0399(.90)	.0147(.88)	.0027(.87)	.0001(.88)
	Bootstrap	.0404(.91)	.0148(.93)	.0028(.87)	.0001(.87)
	Bootstrap-t	.0429(.91)	.0159(.91)	.0029(.92)	.0001(.92)
.05	Asymptotic	.2419(.92)	.0739(.88)	.0151(.88)	.0007(.88)
	Classical	.2001(.90)	.0734(.86)	.0139(.86)	.0007(.87)
	Bootstrap	.2023(.92)	.0738(.84)	.0139(.81)	.0007(.87)
	Bootstrap-t	.2151(.92)	.0793(.91)	.0148(.90)	.0007(.93)
.1	Asymptotic	.4846(.91)	.1480(.87)	.0302(.89)	.0013(.89)
	Classical	.4007(.89)	.1469(.86)	.0278(.88)	.0013(.87)
	Bootstrap	.4052(.91)	.1480(.87)	.0278(.87)	.0013(.87)
	Bootstrap-t	.4316(.91)	.1590(.92)	.0298(.93)	.0014(.91)

Table 8
Different Confidence Intervals
N = 75

σ	Methods	α_1	α_2	β_1	β_2
.01	Asymptotic	.0489(.92)	.0081(.88)	.0025(.91)	4.94E-5(.88)
	Classical	.0404(.93)	.0082(.88)	.0023(.93)	4.95E-5(.88)
	Bootstrap	.0406(.95)	.0081(.88)	.0024(.93)	4.91E-5(.88)
	Bootstrap-t	.0422(.91)	.0085(.91)	.0025(.92)	5.11E-5(.91)
.05	Asymptotic	.2447(.92)	.0408(.89)	.0127(.91)	.0002(.89)
	Classical	.2021(.92)	.0411(.88)	.0118(.93)	.0002(.85)
	Bootstrap	.2030(.92)	.0405(.88)	.0118(.93)	.0002(.88)
	Bootstrap-t	.2114(.92)	.0424(.89)	.0123(.93)	.0003(.91)
.1	Asymptotic	.4900(.91)	.0816(.88)	.0255(.92)	.0005(.88)
	Classical	.4047(.92)	.0822(.88)	.0236(.93)	.0005(.88)
	Bootstrap	.4069(.95)	.0813(.88)	.0237(.93)	.0005(.88)
	Bootstrap-t	.4234(.92)	.0849(.92)	.0248(.93)	.0005(.91)

Table 9
Different Confidence Intervals
N = 100

σ	Methods	α_1	α_2	β_1	β_2
.01	Asymptotic	.0491(.92)	.0056(.88)	.0024(.92)	2.51E-5(.88)
	Classical	.0406(.92)	.0057(.86)	.0022(.89)	2.52E-5(.87)
	Bootstrap	.0411(.91)	.0057(.87)	.0022(.89)	2.49E-5(.87)
	Bootstrap-t	.0422(.91)	.0058(.91)	.0023(.90)	2.58E-5(.90)
.05	Asymptotic	.2459(.92)	.0281(.88)	.0119(.93)	.0001(.88)
	Classical	.2034(.92)	.0284(.84)	.0109(.89)	.0001(.87)
	Bootstrap	.2057(.91)	.0283(.86)	.0109(.88)	.0001(.87)
	Bootstrap-t	.2111(.91)	.0292(.91)	.0114(.90)	.0001(.90)
.1	Asymptotic	.4925(.92)	.0562(.89)	.0239(.92)	.0003(.89)
	Classical	.4072(.92)	.0569(.88)	.0220(.89)	.0003(.88)
	Bootstrap	.4113(.91)	.0565(.88)	.0219(.89)	.0002(.85)
	Bootstrap-t	.4225(.91)	.0583(.90)	.0227(.91)	.0003(.90)